



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

ParCorFull: a Parallel Corpus Annotated with Full Coreference

Citation for published version:

Lapshinova-Koltunski, E, Hardmeier, C & Krielke, P 2018, ParCorFull: a Parallel Corpus Annotated with Full Coreference. in *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*. European Language Resources Association (ELRA), Miyazaki, Japan, pp. 423-428, 11th Edition of the Language Resources and Evaluation Conference, Miyazaki, Japan, 7/05/18. <<https://www.aclweb.org/anthology/L18-1065>>

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Publisher's PDF, also known as Version of record

Published In:

Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



ParCorFull: a Parallel Corpus Annotated with Full Coreference

Ekaterina Lapshinova-Koltunski, Christian Hardmeier, Pauline Krielke

Saarland University, Uppsala University, Saarland University

e.lapshinova@mx.uni-saarland.de, christian.hardmeier@lingfil.uu.se, pauline.krielke@gmail.com

Abstract

In this paper, we describe a parallel corpus annotated with full coreference chains that has been created to address an important problem that machine translation and other multilingual natural language processing (NLP) technologies face – translation of coreference across languages. Recent research in multilingual coreference and automatic pronoun translation has led to important insights into the problem and some promising results. However, its scope has been restricted to pronouns, whereas the phenomenon is not limited to anaphoric pronouns. Our corpus contains parallel texts for the language pair English-German, two major European languages. Despite being typologically very close, these languages still have systemic differences in the realisation of coreference, and thus pose problems for multilingual coreference resolution and machine translation. Our parallel corpus with full annotation of coreference will be a valuable resource with a variety of uses not only for NLP applications, but also for contrastive linguists and researchers in translation studies. This resource supports research on the mechanisms involved in coreference translation in order to develop a better understanding of the phenomenon.

Keywords: coreference, full coreference, cross-lingual coreference resolution, coreference annotation, linguistic annotation, machine translation, multilingual NLP

1. Introduction

We present a parallel corpus containing full coreference annotation that has been created to address an important problem affecting machine translation (MT) and multilingual NLP technologies: translation of coreference across languages. The corpus is available from the LINDAT repository at <http://hdl.handle.net/11372/LRT-2614>.

Texts of various genres often contain recurring references to objects and other discourse entities, realised with a variety of linguistic devices such as noun phrases (NPs), pronouns or other linguistic means. Devices referring to the same entity are said to corefer. The coreference relation is shared across all languages. However, languages differ considerably in the range of linguistic means triggering this relation (Kunz and Steiner, 2012; Kunz and Lapshinova-Koltunski, 2015; Novák and Nedoluzhko, 2015). The choice between these referring expressions is governed by language-specific constraints. Differences in their realisation give rise to transformation patterns used to create coherent translations. In translation, references in the source language (SL) must be rendered with appropriate linguistic devices from the repertoire of the target language (TL), with different constraints. For instance, pronouns and adjectives in German (DE) are subject to grammatical gender agreement, whereas in English (EN), only person pronouns have this marking and adjectives are unmarked.

Recent research in multilingual coreference and automatic pronoun translation has led to important insights into the problem and some promising results, but a working solution for coreference translation in an end-to-end MT has not been demonstrated yet. Research on automatic coreference translation has been restricted to pronouns, but the phenomenon is by no means limited to anaphoric pronouns. Example (1) illustrates a coreference chain expressing a relation of comparison, where we have a closed class of explicit markers for establishing this type of relation in English

(*imaginary ones*). German, however, uses an elliptical noun phrase (*imaginäre [...]*).

- (1) ...I would make an effort to tell them we have real sciences, hard sciences, we don't need [*imaginary ones*]. - ...ich würde mir extra Mühe geben, ihnen zu erzählen, dass wir richtige Wissenschaften haben, hieb- und stichfeste Wissenschaften, wir brauchen [*keine imaginären*].

Moreover, negation in German can be expressed not only with the adverb *nicht*, but also with the indefinite pronoun *kein*. This pronoun changes its form depending on the case, number and gender (*keine* in example (1) is a plural accusative), which also influences the form of the following adjective: *imaginären* (plural accusative). This form is dependent on the antecedent (*Wissenschaften*) of the nominal ellipsis. A translation error such as an incorrect inflection (*imaginäre/imaginäres/imaginärer*) may destroy the coreference chain. Interpreting referential expressions is therefore essential for correct translation.

Even where the systemic options for coreference devices coincide, we can find frequent alternations in the use of demonstrative pronouns in German as in example (2).

- (2) We work for prosperity and opportunity because [*they*]'re right. [*It*]'s the right thing to do. – Wir arbeiten für Wohlstand und Chancen, weil [*das*] richtig ist. Wir tun [*damit*] das Richtige ('We work for prosperity and opportunity because [*that*] is right. We do [*thereby*] the right').

The English example uses the personal pronouns *they* and *it* to refer to the entities *prosperity* and *opportunity* in the first case, and to the event *working for prosperity and opportunity* in the second. The example translation from a parallel corpus uses the demonstrative *das* and the demonstrative deictic *damit*, referring to the event *working for prosperity and opportunity* in both cases, but encoding an additional

logico-semantic relation of instrumentality in the second. This is one of the typical cases of translation between English and German where the coreference relation as such is preserved, but it is not coreference between exactly the same entities in both cases, and it is semantically enriched by the instrumental relation in the second. An MT system is very likely to output the personal pronoun *sie* instead of *das* (*Wir arbeiten für Wohlstand und Chancen, weil [sie] richtig sind*) making this pronoun refer to the entities and not the event, which would sound less natural in German.

2. Related Work

The challenge of translating pronouns has been a recurring topic in recent studies. There are a few corpus-based studies of coreference translation (Novák and Nedoluzhko, 2015; Novák et al., 2013; Guillou and Webber, 2015). For the languages under analysis, it has been empirically shown to be a relevant problem (Hardmeier and Federico, 2010; Guillou, 2016). In the MT community, the awareness of the problem has been increased with three recent shared tasks on pronoun translation (Hardmeier et al., 2015; Guillou et al., 2016; Loáiciga et al., 2017). In recognition of the difficulty of the problem, test suite-based evaluation methods for pronoun translation have been proposed (Guillou and Hardmeier, 2016; Bawden et al., 2017).

At the same time, coreference translation and multilingual coreference resolution is still a complex problem, as we observe a widespread lack of understanding of this phenomenon. Existing coreference resolution tools are known to be unreliable as they introduce an unacceptable number of errors, and therefore manually annotated parallel resources are absolutely indispensable for the development of coreference-aware MT systems and other multilingual language technologies, including cross-lingual coreference resolution (Grishina, 2017; Novák and Žabokrtský, 2014; Green et al., 2011), information extraction (Lee et al., 2012; Zelenko et al., 2004) and question answering (Morton, 1999; Hartrumpf et al., 2008). Most existing coreference corpora are not parallel. The only resources for the language pair English-German that are known to us include the GECCo corpus (Lapshinova-Koltunski and Kunz, 2014), the ParCor corpus (Guillou et al., 2014) and the multilingual coreference corpus described by (Grishina and Stede, 2015). The first corpus contains annotations of the source texts only and is available with restrictions on some texts. The second resource considers only pairwise annotation of anaphoric pronouns and their antecedents. The third corpus, although containing annotations of all referring expressions appearing in a coreference chain, is very small (ca. 11,000 words per language).

For this reason, we have created an English-German parallel corpus which contains annotation of full coreference chains on the basis of the ParCor corpus. The annotation scheme takes inspiration from the schemes used in all the three resources mentioned above (Kunz, 2012; Guillou et al., 2014; Grishina and Stede, 2016). In contrast to other existing coreference schemes that were designed for monolingual datasets, these were elaborated for a multilingual corpus and will allow us to obtain uniform nominal coreference annotations which facilitates extension to further

languages in the future.

3. Annotation Categories

A detailed description of categories and disambiguation rules are needed to guarantee consistency throughout the whole process of annotation. Our annotation guidelines are based on the three existing ones described by Grishina and Stede (2016), Guillou et al. (2014) and Kunz (2012). They address the segmentation of nominal elements, the annotation of different antecedent and anaphora types and examples of various problematic cases (Lapshinova-Koltunski and Hardmeier, 2017).

Segmentation Annotated elements (markables) include: Pronouns, nouns, nominal phrases or elliptical constructions that are parts of a coreference pair (antecedent-anaphora), as well as verbal phrases or clauses being antecedents of event anaphora.

Types of antecedents In our framework, we define two different types of antecedents: entities and events. Entities can either be represented by a pronoun or an NP. Events can be represented by a VP as in (3-a), a clause as in (3-b) or a set of clauses. Antecedents can be split as in (3-b) (multiple elements "prosperity and opportunity" constitute one antecedent – all components of the antecedent are linked to the referring expression "they"). If there is no explicit antecedent (in some cases, a referring expression is anaphoric, but no specific antecedent can be found in the text) the position of the antecedent is left open.

- (3) a. ... you have to basically [combine everything you learned from project one and project two]. ultimately [that]'s the goal .
- b. [We work for [prosperity] [and opportunity]] because [they]'re right. [It]'s the right thing to do.

Types of anaphora We include two types of referring expressions (anaphors) into our analysis: Pronouns and nominal phrases. Coreferring **pronouns** include personal, demonstrative, relative and reflexive pronouns. Note that demonstrative pronouns may also refer to locations (*there, here*) and time (*then, now*). We also include pronominal adverbs in the category of demonstrative pronouns. Pronominal adverbs are formed by replacing a preposition and a pronoun, like *gegen+das* → *dagegen* in example (4). They exist in both English and German, but are used differently. In German, they are very common, but in English, they sound rather archaic and are generally avoided. adverbs are not considered in most coreference annotation schemes. However, they constitute around 8% of all referring expressions in the German language¹ and are especially frequent in spoken and spoken-like language.

- (4) Viele Amerikaner haben Probleme mit [Rassismus]; doch wir sind [dagegen] immun.

Coreferring **nominal phrases** include proper names (*Herr Almeida Freire* in example (5-a)), nominal premodifiers as

¹This number is based on the annotations available in the German part of the GECCo corpus (Lapshinova-Koltunski and Kunz, 2014).

in (5-b), full nominal phrases (used with a definite article or a demonstrative modifier as in example (5-c)) and nominal phrases with quantifiers (*all people* in the meaning *all these people*). Generic nouns can co-refer with definite full NPs or pronouns, but not with other generic nouns, see (5-d).

- (5) a. *In [seiner] EWSA-Stellungnahme zum "Bericht der Kommission zur Beobachtung des Handelsmarktes" schreibt [Herr Almeida Freire]...*
- b. *The unionists used to be [[EU] supporters], but now they are questioning how [it] has developed...*
- c. *This past spring, the U.S. Department of Education issued [a report, The Condition of Education 2000]. [The report] found that...*
- d. *[Computers] are expensive. But [they] are useful. Computers cost a lot of money.*

As shown in example (1) above, linguistic chains may also include *substitution* and *ellipsis* in addition to referring expressions. These trigger a type reference relation (as opposed to a relation of identity) between referents belonging to the same class (Kunz and Steiner, 2013; De Beaugrande and Dressler, 1981). In substitution patterns, the referring expression is replaced with another element (example (6-a)). In ellipsis, it is completely left out, and the reference is implicit (example (6-b)).

- (6) a. *Do you prefer the blue shirt or [the red shirt]? – I would like the red [one].*
- b. *...if I take any one of these balls... and I count how many [neighboring balls] that there are around it, the answer's always twelve [].*

Substitution and ellipsis are mostly analysed within separate chains in other studies. We include them into our framework, since they often occur in similar contexts as coreference if considered cross-lingually, as again was shown in example (1). In our framework, substitution and ellipsis are subdivided into their structural types, according to the omitted/substituted element: nominal ellipsis (7-a), verbal substitution (7-b) and clausal substitution in (7-c).

- (7) a. *You might have to come up afterwards to count but if I take any one of these balls in the middle and I count how many [neighboring balls] that there are around it, the answer's always twelve [].*
- b. *You'll see that it had [to accommodate] an incredible range of functions much more elaborate than any temple or palace in the past would have [done].*
- c. *[Does everybody have a handout, for today]? If [not] Aaron's got handouts.*
- d. *[So, well, any more questions]? – [no], okay, ...*
- e. *[How many slices do you want]? – "[Two]", I said.*

Following Menzel (2017), we also define two additional classes for ellipsis: yes-no type as in (7-d) and mixed type

(a combination of nominal and verbal or clausal) as illustrated in (7-e).

Another category that is considered here but is excluded from most analyses is that of comparative reference, which does not trigger co-reference in the strict sense. Together with other cases (substitution and ellipsis) it instead involves type reference, co-classification or "sloppy identity", see (Kunz and Steiner, 2012). The linguistic means signaling comparative reference include such words as *same*, *equal*, *identical* or particular adjectives in the comparative form. We distinguish between general and particular comparison, the first referring to a general relation of comparison between two entities (8-a) and the latter referring to particular comparative features of two entities (8-b).

- (8) a. *So what do you think happened to [these design students] ? [...] We did the exercise again with [the same students] .*
- b. *That car over there is very [fast] . But well, my uncle drives an even [faster] one.*

4. Annotation Process

The annotations were created with the annotation tool MMAX2, including all of the above mentioned categories. The annotation scheme created for this task allows the annotator to define each markable as a certain mention type (pronoun, NP, VP or clause). The mentions can be defined further in terms of their cohesive function (antecedent, anaphoric, cataphoric, comparative, substitution, ellipsis, extratextual, pleonastic-it, apposition). Antecedents can either be annotated as simple or split, and as entity or event. For anaphoric expressions the scheme includes singular/plural agreement with the antecedent and subject/non-subject position of the expression. The annotation scheme also covers pronoun type (personal, possessive, demonstrative, reflexive, relative) and modifier types of NPs (possessive, demonstrative, definite article, or none for proper names). An example of the MMAX2 interface with a visualisation of a coreference chain is illustrated in Figure 1.

All annotations were performed by highly experienced well-trained annotators with linguistic background in order to ensure maximum accuracy.

5. Data Selection

We used existing resources and extended them with: (1) complete annotation of full coreference chains; (2) additional referring expressions to achieve full coreference chains. The resources we are using as a basis include the ParCorpus (Guillou et al., 2014) and the dataset used for the DiscoMT workshop shared task (Hardmeier et al., 2015). To increase the variety in register and genre, we included some additional data taken from the test sets of the news translation shared task at the Conference on Machine Translation (Bojar et al., 2017, WMT17). Table 1 provides an overview of the total number of tokens included into our corpus.

For the DiscoMT dataset, the existing annotations covered only English. We extended the existing dataset into a parallel corpus by adding the corresponding translation from

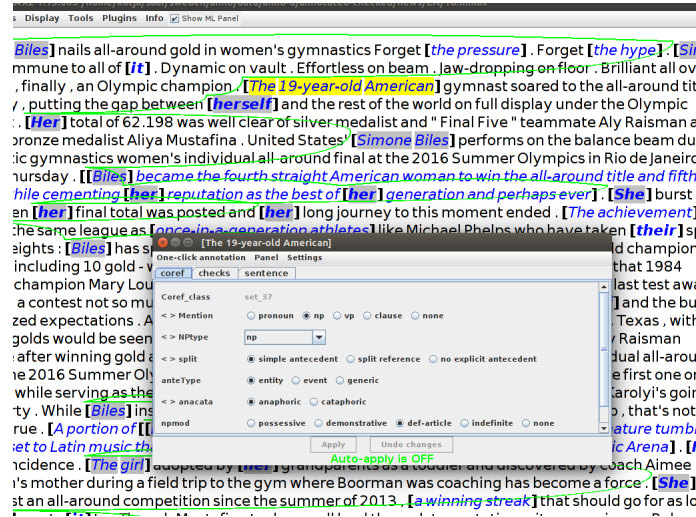


Figure 1: A coreference chain visualised in MMAX2

language	ParCor	DiscoMT	WMT news	total
English	31,971	39,764	10,644	82,379
German	30,305	37,452	10,593	78,350
total	62,276	77,216	21,237	160,729

Table 1: Corpus data

English into German. Overall, we completed the annotations by adding all types of referring expressions 71,735 tokens of the English data and 30,305 of the German data. Around 48,000 tokens of data in German (translations of the English TED talks contained in the DiscoMT data and the news texts from the WMT data), as well as 10,644 tokens of the English data did not contain any annotations and were thus annotated from scratch. The total number of tokens in the annotated corpus amounts to ca. 160,000.

The annotated resource that we have created represents a reasonably-sized data set for training coreference resolution components that can be used for MT or other cross-lingual applications. It is comparable in size (with a larger amount of text, but fewer annotated mentions) to the AR-RAU corpus (Poesio and Artstein, 2008), which features a similarly rich coreference annotation and covers a greater variety of genres, but does not include multilingual parallel text. Although the amount of data is not enough to train an MT system, this dataset will be large enough for MT tuning, testing and evaluation, which is an important improvement over the existing data situation.

6. Annotation Results

We present an overview of the annotated structures (absolute numbers) in Table 2 below.

In total, the corpus contains about 15,000 annotated mentions at the moment. The annotated mentions are classified according to their morpho-syntactic type: pronouns (pronoun), nominal phrases (np), verbal phrases (vp) and clauses (clause). This differentiation was introduced for a practical reason, as it permits classifying mentions further according to their function or the role in a coreference

	English	German	total
pronoun	4,650	4,269	8,919
np	2,485	2,611	5,096
vp	133	132	265
clause	335	312	647
total mentions	7,603	7,324	14,927

Table 2: Annotated mentions and their subcategories

chain. As seen from the table, German texts contain more markables, i.e. more referring expressions.

	English	German	total
number of chains	2,319	2,425	4,744
average chain length	2.94	2.81	2.87

Table 3: Annotated chains

The number of full coreference chains in the data amounts to 4,744 (see Table 3). We also calculate the average chain length (total number of mentions/total number of chains). The German translations contain more chains than their English sources, but on average, these chains are shorter. To evaluate the reliability of the annotated coreference chains, we created a second annotation of two files in each language. The inter-annotator dataset included TED talks 785 and 790 and was composed of 6,253 English and 5,975 German tokens. As a measure of inter-annotator agreement, we computed the mention overlap and entity-based CEAF scores (Luo, 2005) between the two annotations, treating our regular annotator as the hypothesis to be evaluated and

the second annotator as the reference. The scores were calculated with the CoNLL reference scorer implementation (Pradhan et al., 2014) and are shown in Table 4 as a macro-average over the two documents.

	Precision	Recall	F-score
English			
mentions	89.20%	73.89%	80.71%
CEAF _e	82.90%	67.13%	74.13%
German			
mentions	84.80%	69.76%	76.54%
CEAF _e	72.53%	60.36%	65.88%

Table 4: Inter-annotator metrics for coreference chains

As seen from the table, we observe a better agreement for the English texts. We suppose that the reason for the greater disagreement for German texts is the complexity of the linguistic structures triggering coreference in this language. However, a more detailed analysis of the agreement results is needed to understand the reasons. We plan to do this in future work.

We also performed automatic inconsistency checks to prove if the annotated data contains any (1) marked mentions outside of chains; (2) antecedents of chains that are not marked as first elements of chains; (3) some other error types. The detected errors were then corrected by the annotators. Besides that, the annotator added the following categories (that were not included into the annotation scheme at the very beginning): (a) bare nouns; (b) indefinite nouns and (c) quantifiers (both) as demonstratives.

7. Conclusion and Future Work

The differences in coreference realisation in multiple languages present a huge challenge to machine translation and are of interest for contrastive linguists and researchers in translation studies. A parallel corpus with full annotation of coreference is a valuable resource with a variety of uses. The corpus will help us study the mechanisms involved in coreference translation in order to develop a better understanding of the phenomenon. It will serve as a resource for creating and evaluating coreference-aware MT systems without having to rely on notoriously inaccurate automatic coreference resolvers. Finally, it can also be used as a training and development resource for the creation of multilingual or monolingual coreference resolution systems. Moreover, we address the demand for better approaches to evaluate complex linguistic phenomena that are not covered by existing annotation schemes.

8. Acknowledgements

The annotation work was funded by the European Association for Machine Translation. Christian Hardmeier was supported by the Swedish Research Council under grants 2012-916 and 2017-930.

9. Bibliographical References

- Bawden, R., Sennrich, R., Birch, A., and Haddow, B. (2017). Evaluating discourse phenomena in neural machine translation. *ArXiv e-prints*, 1711.00513.
- Bojar, O., Chatterjee, R., Federmann, C., Graham, Y., Haddow, B., Huang, S., Huck, M., Koehn, P., Liu, Q., Logacheva, V., Monz, C., Negri, M., Post, M., Rubino, R., Specia, L., and Turchi, M. (2017). Findings of the 2017 conference on machine translation (wmt17). In *Proceedings of the Second Conference on Machine Translation, Volume 2: Shared Task Papers*, pages 169–214, Copenhagen, Denmark, September. Association for Computational Linguistics.
- De Beaugrande, R.-A. and Dressler, W. U. (1981). *Einführung in die Textlinguistik*. Niemeyer, Tübingen.
- Green, S., Andrews, N., Gormley, M. R., Dredze, M., and Manning, C. D. (2011). Cross-lingual coreference resolution: A new task for multilingual comparable corpora. Technical Report 6, HLTCOE, Johns Hopkins University.
- Grishina, Y. and Stede, M., (2016). *Parallel coreference annotation guidelines*.
- Grishina, Y. (2017). Combining the output of two coreference resolution systems for two source languages to improve annotation projection. In *Proceedings of the Third Workshop on Discourse in Machine Translation (DiscoMT), EMNLP 2017*, Copenhagen, Denmark.
- Guillou, L. and Hardmeier, C. (2016). PROTEST: A test suite for evaluating pronouns in machine translation. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016)*, pages 636–643, Portorož, Slovenia. European Language Resources Association (ELRA).
- Guillou, L. and Webber, B. (2015). Analysing ParCor and its translations by state-of-the-art SMT systems. In *Proceedings of the Second Workshop on Discourse in Machine Translation (DiscoMT), EMNLP 2015*, Lisbon, Portugal.
- Guillou, L., Hardmeier, C., Smith, A., Tiedemann, J., and Webber, B., (2014). *ParCor 1.0: Pronoun Coreference Annotation Guidelines*. Edinburgh, Uppsala.
- Guillou, L., Hardmeier, C., Nakov, P., Stymne, S., Tiedemann, J., Versley, Y., Cettolo, M., Webber, B., and Popescu-Belis, A. (2016). Findings of the 2016 WMT shared task on cross-lingual pronoun prediction. In *Proceedings of the First Conference on Machine Translation (WMT16)*, Berlin, Germany.
- Guillou, L. (2016). *Incorporating Pronoun Function into Statistical Machine Translation*. Ph.D. thesis, School of Informatics. University of Edinburgh.
- Hardmeier, C. and Federico, M. (2010). Modelling pronominal anaphora in statistical machine translation. In *Proceedings of the IWSLT (International Workshop on Spoken Language Translation)*, pages 283–289, Paris, France.
- Hardmeier, C., P., Nakov, S., Stymne, J., Tiedemann, Y., Versley, and Cettolo, M. (2015). Pronoun-focused MT and cross-lingual pronoun prediction: Findings of the 2015 DiscoMT shared task on pronoun translation. In

- Proceedings of the Second Workshop on Discourse in Machine Translation (DiscoMT)*, EMNLP 2015, pages 1–16, Lisbon, Portugal.
- Hartrumpf, S., Glöckner, I., and Leveling, J. (2008). Coreference resolution for questions and answer merging by validation. In Carol Peters, et al., editors, *Advances in Multilingual and Multimodal Information Retrieval: 8th Workshop of the Cross-Language Evaluation Forum, CLEF-200, Budapest, Hungary, September 19-21, 2007*, pages 269–272. Springer, Berlin, Heidelberg.
- Kunz, K. and Lapshinova-Koltunski, E. (2015). Cross-linguistic analysis of discourse variation across registers. *Special Issue of Nordic Journal of English Studies*, 14(1):258–288.
- Kunz, K. and Steiner, E. (2012). Towards a comparison of cohesive reference in English and German: System and text. In M. Taboada, et al., editors, *Contrastive Discourse Analysis. Functional and Corpus Perspectives*. Equinox, London.
- Kunz, K. and Steiner, E. (2013). Cohesive substitution in English and German: A contrastive and corpus-based perspective. In Karin Aijmer et al., editors, *Advances in Corpus-Based Contrastive Linguistics. Studies in honour of Stig Johansson*, pages 201–232. John Benjamins, Amsterdam.
- Kunz, K., (2012). *Richtlinien für die Korrektur von kohäsiven Referenzmitteln*.
- Lapshinova-Koltunski, E. and Hardmeier, C., (2017). *Coreference Corpus Annotation Guidelines*, December.
- Lee, H., Recasens, M., Chang, A., Surdeanu, M., and Jurafsky, D. (2012). Joint entity and event coreference resolution across documents. In *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, EMNLP-CoNLL '12, pages 489–500, Jeju Island, Korea.
- Loáiciga, S., Stymne, S., Nakov, P., Hardmeier, C., Tiedemann, J., Cettolo, M., and Versley, Y. (2017). Findings of the 2017 DiscoMT shared task on cross-lingual pronoun prediction. In *Proceedings of the Third Workshop on Discourse in Machine Translation (DiscoMT)*, EMNLP 2017, pages 1–16, Copenhagen, Denmark.
- Luo, X. (2005). On coreference resolution performance metrics. In *Proceedings of Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing*, pages 25–32, Vancouver, British Columbia, Canada.
- Menzel, K. (2017). *Understanding English-German contrasts: a corpus-based comparative analysis of ellipses as cohesive devices*. Ph.D. thesis, Universität des Saarlandes, Saarbrücken.
- Morton, T. S. (1999). Using coreference for question answering. In *Proceedings of the Workshop on Coreference and its Applications*, CorefApp-99, pages 85–89, College Park, Maryland.
- Novák, M. and Nedoluzhko, A. (2015). Correspondences between Czech and English coreferential expressions. *Discours*, 16.
- Novák, M. and Žabokrtský, Z. (2014). Cross-lingual coreference resolution of pronouns. In *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers*, pages 14–24, Dublin, Ireland.
- Novák, M., Žabokrtský, Z., and Nedoluzhko, A. (2013). Two case studies on translating pronouns in a deep syntax framework. In *Proceedings of the 6th International Joint Conference on Natural Language Processing*, pages 1037–1041, Nagoya, Japan.
- Pradhan, S., Luo, X., Recasens, M., Hovy, E., Ng, V., and Strube, M. (2014). Scoring coreference partitions of predicted mentions: A reference implementation. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 30–35, Baltimore, Maryland.
- Zelenko, D., Aone, C., and Tibbetts, J. (2004). Coreference resolution for information extraction. In *Proceedings of the Conference on Reference Resolution and Its Applications*.

10. Language Resource References

- Grishina, Y. and Stede, M. (2015). Knowledge-lean projection of coreference chains across languages. In *Proceedings of the 8th Workshop on Building and Using Comparable Corpora*, page 14, Beijing, China.
- Guillou, L., Hardmeier, C., Smith, A., Tiedemann, J., and Webber, B. (2014). ParCor 1.0: A parallel pronoun-coreference corpus to support statistical MT. In *Proceedings of the 9th International Conference on Language Resources and Evaluation (LREC-2014)*, Reykjavik, Iceland.
- Lapshinova-Koltunski, E. and Kunz, K. (2014). Annotating cohesion for multilingual analysis. In *Proceedings of the 10th Joint ACL - ISO Workshop on Interoperable Semantic Annotation*, pages 57–64, Reykjavik, Iceland.
- Poesio, M. and Artstein, R. (2008). Anaphoric annotation in the ARRAU corpus. In Bente Maegaard Joseph Mariani Jan Odijk Stelios Piperidis Daniel Tapias Nicoletta Calzolari (Conference Chair), Khalid Choukri, editor, *Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC'08)*, Marrakech, Morocco.